

University of Groningen

Putting Popper to Work

Derksen, Maarten

Published in:
Theory & Psychology

DOI:
[10.1177/0959354319838343](https://doi.org/10.1177/0959354319838343)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Derksen, M. (2019). Putting Popper to Work. *Theory & Psychology*, 29(4), 449-465.
<https://doi.org/10.1177/0959354319838343>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Putting Popper to work

Maarten Derksen

University of Groningen, Netherlands

Theory & Psychology
2019, Vol. 29(4) 449–465
© The Author(s) 2019Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0959354319838343
journals.sagepub.com/home/tap**Abstract**

In response to what are seen as fundamental problems in Psychology, a reform movement has emerged that finds inspiration in philosophy of science, the work of Karl Popper in particular. The reformers attempt to put Popper into practice and create a discipline based on the principles of critical rationalism. In this article I describe the concrete sociotechnical practices by which the reformers attempt to realise their ideals, and I argue that they go a long way towards bridging the gap between rules and practice that sociologists of science Mulkay and Gilbert had identified in their study of the role of Popper's philosophy in the work of scientists. Second, I note the considerable resistance that the reformers meet and the disruptive force of their work. I argue that this disruption is productive and raises fundamental questions regarding psychology and its object of study.

Keywords

crisis, critical rationalism, falsification, reform, replication, social media

It is a feature of the reform movement¹ that is sweeping through psychology at the moment that it is explicitly inspired by philosophy of science, the work of Karl Popper in particular. References to Popper and like-minded philosophers of science such as Paul Meehl are used to support arguments about what science is, what has caused the current problems, and how they should be solved.² The reform movement is an epistemic project that is informed by epistemology. This would have pleased Popper, who wrote, in the preface to the English edition of *The Logic of Scientific Discovery*, that the theory of knowledge, including his own, aims to “enable us not only to know more about knowledge, but also to contribute to the advance of knowledge – of scientific knowledge that is” (1972/2002a, p. xxii).

This development may amuse those who, like me, have taken on board the lessons of the empirical turn in the study of science. One of those lessons after all appeared to be that science does not work the way Karl Popper thought it should. When sociologists of science in the 1970s entered laboratories to study the actual process of science they concluded that it doesn't follow the rules that Popper, or any other philosopher of science for that matter,

Corresponding author:Maarten Derksen, University of Groningen, Grote Kruisstraat 2/1, Groningen 9712 TS, Netherlands.
Email: m.derksen@rug.nl

had laid down. Research is messy, researchers are motivated by more than a desire for objective truth, and facts are not discovered but constructed in a process that involves many more actors than those allowed by traditional philosophers of science. When Bruno Latour was asked (in an interview on Dutch television) whether science proceeds by falsification, he laughed derisively and answered: "That is the textbook version of science, I don't think there is any case where this works as nicely as this" (Schepens, 2008, at 14:15). However, as much as we post-Kuhnian students of science might be tempted to dismiss the psychology reformers as naive, our first step should be to study this development empirically, including the role Popper's ideas play in it.

Michael Mulkay and Nigel Gilbert were here before us. In their article "Putting Philosophy to Work: Karl Popper's Influence on Scientific Practice" (1981) they reported on interviews they had conducted with biochemists in the context of their study of a controversy in that field (see also N. Gilbert & Mulkay, 1984). These researchers regularly mentioned Popper, even though they never referred to him in their primary literature. What was the role of Popper's ideas in their own thinking and research? What work did Popper do for them? It became clear from the interviews that the rules that Popper had formulated did not function to constrain action. Even the interviewees who identified as Popperians acknowledged they did not look to Popper's methodological rules as prescriptions for their day to day laboratory work. Rather, Popper was used as an evaluative resource, a way to judge "good" and "bad" after the fact. Popper's philosophy of science was used to describe what was good about a particular study or researcher, but not to determine how to proceed in doing research.

Mulkay and Gilbert saw a fundamental truth about rules reflected here: the relationship between rules and practice is "essentially indeterminate" (Mulkay & Gilbert, 1981, p. 404). Rules do not determine their own application. Following a rule always requires an interpretation of what the rule means in this particular situation. In science, the gap between rules and action is particularly wide because at the forefront of research, novel situations are being created: new techniques, new instruments, and of course new phenomena and effects, which all raise the question of how the rules apply in these novel circumstances. For example, Popper had argued that one cannot verify a theory, one can only disprove it. Thus, falsifiability is the mark of a scientific theory, and science should proceed by attempts at falsification. That is clear enough. However, Mulkay and Gilbert noted that whether or not a particular experimental result is a falsification depends on a technical, scientific appraisal of the experiment. In situations of scientific uncertainty, in new lines of research, such appraisals will vary between researchers. "Consequently, when there is uncertainty, the Popperian rules cannot provide a straightforward guide for scientists' actions or decisions. There is a gap between rule and particular action which can only be bridged by the very scientific choice which the rule is intended to constrain" (Mulkay & Gilbert, 1981, p. 398).

Popper's work does not provide any guidance for how to deal with this interpretative challenge. His rules of method were based on the rational reconstruction of scientific achievements rather than on a description of actual scientific practice. From this hindsight perspective the interpretative work that was required in scientific practice becomes invisible. As a result this (or indeed any) prescriptive philosophy of science on its own can give little guidance for this work. Popper "remains unclear about the connection between the

formal analysis of scientific belief systems and the provision of rules of action; and ... he hasn't considered in detail how his rules of scientific method are to be put into practical effect" (Mulkay & Gilbert, 1981, p. 392). To tighten the link between prescription and action, Mulkay and Gilbert argue, the rules must become embodied in a social practice, "so that potential actors have access to a corpus of exemplary instances, they are guided in their efforts by skilled interpreters, and they are subject to various kinds of direct control" (Mulkay & Gilbert, 1981, p. 407).³ Rules will never determine action unambiguously, but they can be made more effective through the interpretative work of a community of researchers, who translate general methodological rules (such as "always expose your theories to the possibility of refutation") into more specific directions for what should be done with regard to this specific hypothesis and others of its kind, who negotiate difficult cases, draw up guidelines, and sanction those who contravene them.

The same question that Mulkay and Gilbert asked 38 years ago is relevant today with regard to the current reformers in psychology: how do these psychologists put Popper to work? The history of science studies following Mulkay and Gilbert's paper has given this question added poignancy. In 1981 Popper's influence was still great (and he was still alive), and that generation of sociologists of science partly defined itself in relation to his work: by turning Popper on his head, for example, as in this paper, and investigating how norms work in practice. Since then Popper has gradually receded from view in science studies, and is no longer relevant, even as a foil. What then should we make of this apparent renaissance of Popperian thinking in psychology?

Philosophy in practice

Ever since the reform movement in psychology began to coalesce in 2011, replicability has been its main concern.⁴ The frequency with which even high-profile studies fail to replicate is seen as an indication of fundamental problems in the usual research practices of the discipline. A lack of transparency is seen as a major cause of the problems: a lack of disclosure about the actual research process and its results allows researchers to present their studies in a favourable but misleading way and keeps null results ("failed experiments") in the file drawer. As a result, the discipline's archive may look like a corpus of scientific success stories, but it is actually "a vast graveyard of undead theories" (Ferguson & Heene, 2012).

In the proposals for improvement that have been appearing regularly over the past six years, statements about "what science is" have an important place.⁵ In contrast to Mulkay and Gilbert's biochemists,⁶ quite a number of these reformers in psychology read philosophy of science. For instance, a recent blog post in which the author argued that null hypothesis significance testing is compatible with Popper's ideas about falsification mentioned Duhem, Lakatos, Laudan, Van Fraassen, and Feyerabend as well as Popper (Lakens, 2017a). Another example: Zoltán Dienes' *Understanding Psychology as a Science* (2008), a textbook that leans heavily on Popper, is enthusiastically recommended as summer reading (Lakens, 2017b; Srivastava, 2017). Even when Popper is not mentioned, science is depicted in a way that largely conforms to his philosophy of science, in that falsifiability, falsification, and replication are seen as crucial elements of the scientific process. Reformers emphasise that scientific theories must be falsifiable. This is

mostly presented as self-evident, but sometimes Popper is referred to (e.g., LeBel & Peters, 2011, p. 373). Falsification, the reformers believe, “is achieved via meticulously executed series of *direct replications*” (LeBel, 2017, line 8), that is to say, by following the procedure of the original experiment as closely as possible. In this context, the reformers like to quote from section 8 of Popper’s *Logic of Scientific Discovery*, where he states that observations are only inter-subjectively testable when they can be repeated by following specific instructions. The upshot is, in the words of one prominent reformer,

that (1) scientists should replicate their own experiments; (2) scientists should be able to instruct other experts how to reproduce their experiments and get the same results; and (3) establishing the reproducibility of experiments (“direct replication” in the parlance of our times) is a necessary precursor for all the other things you do to construct and test theories. (Srivastava, 2014b, para. 1)

Whereas Popper was primarily an evaluative resource for Mulkay and Gilbert’s biochemists, these psychologists translate Popper’s philosophy of science into a program of reform, with concrete rules of practice, an infrastructure for that practice, and research projects that realise the ideal. The rules and requirements focus, first of all, on restraining the so-called “researcher degrees of freedom,” a term coined by Simmons, Nelson, and Simonsohn (2011) for the leeway that researchers have in the decisions they take for the collection and analysis of their data. Hypothesis testing requires such decisions—sample size, exclusion of outliers, which comparisons to make, etcetera—to be taken before data collection begins, but it is common practice, Simmons et al. noted, to explore various possibilities during data analysis, see which combination produces statistical significance, and only report that result. Another problematic practice, not discussed by Simmons et al. in their paper, is to fit the hypothesis to the data and create a significant result that way, a process known as HARKing: hypothesising after results are known (Kerr, 1998). In all these cases, one is not testing (i.e., attempting to falsify) hypotheses, but generating hypotheses from the data.

The most commonly proposed solution to this problem is pre-registration, where researchers create a detailed plan for data collection and analysis and upload it to a repository, where it gets a date stamp (see, e.g., Bishop, 2013; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). Once the study is completed, other researchers can consult the repository to see whether the study actually followed the registered research plan. Of course, this system can be gamed by registering multiple plans, but that would require intentional bad faith rather than sloppy or questionable research practice.

A related initiative is the registered report (RR), a publication option devised by Chris Chambers (2013), editor of *Cortex*, and now offered by an increasing number of journals in psychology and other disciplines. In an RR, a researcher submits a proposal for a study to a journal, including a detailed plan for data collection and analysis. Reviewers then look at the relevance and importance of the question and the quality of the research plan, and if the proposal is accepted, publication of the final report is guaranteed regardless of the outcome of the study, provided the research plan has been followed. The RR procedure combines the advantage of pre-registration with the guarantee that null results will also be published, thus increasing the chance that theories are

(recognised to have been) falsified. Moreover, registered reports are often offered as a publication option for direct replication studies (RRR, registered replication reports).

Pre-registration and registered reports are presented as enforcing a distinction between “exploratory” and “confirmatory” research, between generating hypotheses and testing them. Both are important, it is emphasised, but testing has to be separate from exploration to properly count as such (e.g., Bishop, 2013; Chambers, 2017a; Neuroskeptic, 2013). Thus, pre-registration and registered reports give an administrative guarantee that what is presented as the test of a prediction was indeed an attempt to falsify a hypothesis and not actually an inductive process of developing theory from data. In contrast to Popper, these reformers explicitly give induction a place in the scientific process, but on condition that it remains strictly separate from “purely confirmatory research” (Wagenmakers et al., 2012). It’s clear moreover that whereas exploration is acknowledged as important, hypothesis testing is considered necessary: without it there is no science.

Social engineering

To Popper, the rules of method that he proposed in *The Logic of Scientific Discovery* (1972/2002a) were the application to science of a more general conception of reason that he called critical rationalism. In *The Logic of Scientific Discovery* and even more in *The Open Society and its Enemies* (2002b), Popper emphasised that scientific objectivity is not the result of the attitude or efforts of individual scientists, but rather the “product of the social or public character of the scientific method” (2002b, p. 491). The rational route to objectivity is openness and “friendly-hostile co-operation” (Popper, 2002b, p. 489). In line with this ideal, the reform movement pursues transparency in all stages of the research process and it has a penchant for vigorous discussion. Equally in line with Popper’s recommendations, it tries to institutionalise these values. Critical rationalism, Popper wrote, implies the necessity for creating institutions that protect “freedom of criticism, freedom of thought” (Popper, 2002b, p. 511). It “establishes something like a moral obligation” to engage in “practical social engineering” (p. 511). Although *The Open Society and its Enemies* (Popper, 2002b) is seldom referred to, the reform movement is doing exactly what Popper thought was necessary: not only does it make enthusiastic use of social media to conduct its “friendly-hostile” discussions, it is also engineering its own social-technological infrastructure to enable transparency, collaboration, and replication.

A particularly successful initiative is the Open Science Framework (OSF, <https://osf.io/>), an online platform developed and maintained by the Center for Open Science (COS), founded by social psychologists Brian Nosek and Jeffrey Spies. The OSF facilitates the kind of open, reproducible science that the reformers advocate. Via the OSF, researchers can pre-register their studies, share materials, store and share data, and collaborate with others. The goal is to make the entire research process transparent, and thus also eminently reproducible. The OSF has hosted a number of replication projects that were started by reformers to put their ideas about replication and falsification into practice. The best known of these was a mammoth collaborative effort to replicate 100 studies from one volume of three psychological journals. Working together in the “Open Science Collaboration” were 269 researchers, mainly from North America and Europe. The

results were alarming: according to the team's assessment, only 39 replication attempts succeeded (Open Science Collaboration, 2015).⁷ The OSF has also hosted several so-called "Many Labs" projects, in which a small number of original studies are each replicated by a number of research teams, to see whether replication success depends on contextual factors such as the nationality of the participants (Klein et al., 2014) or the semester in which students participate in studies (Ebersole et al., 2016). Neither factor appeared to have much influence on whether the original result could be reproduced. Thus, the Many Labs projects had produced evidence against contextual moderators as explanations for replication failure (Srivastava, 2015).

The internet is the reform movement's primary habitat. Blogs are popular outlets for ideas and commentary; blog posts are typically announced on Twitter and Facebook, and then further discussed there, often leading to new blog posts. Web-based platforms like the OSF not only facilitate collaboration but also serve to share preprints (<https://osf.io/preprints/psyarxiv>), following the example of the successful physics preprint repository arXiv. Such online platforms and social media allow fast, free dissemination of ideas, results, and papers, followed by virtually instant, open discussion and critique, often leading to follow-up collaborative projects to test new hypotheses. Bobbie Spellman has argued that replication studies could only become such an important factor in the crisis because news about them spreads quickly over the internet, whereas before, one would typically learn about a failure to replicate, if at all, in a "fortuitous late night conference conversation" (Spellman, 2015, p. 888). The online world, moreover, is one where traditional gatekeepers such as editors and reviewers are much less powerful because anyone can be an editor—publish their own blog, for example—and a reviewer—for instance on PubPeer, the "online journal club" where anyone can post comments on any scientific article (<https://pubpeer.com>). The social practice characterised by transparency and mutual criticism which the reform movement is creating embodies Popper's critical rationalism. A kind of Popperian Open Society is in the making, an Open Psychology, where everything may be subjected to criticism by anyone, and tradition and reputation hold no sway.

Controversy

This budding community of practice is not without its detractors, however. Open Psychology is resisted by a number of social psychologists in particular, who object to what they consider unfair criticism of the status quo, to the way this criticism is expressed, and to the emphasis on direct replication as a *sine qua non* of science.⁸ The discussion between reformers and counter-reformers can become quite heated, as was the case in the controversy in 2014 over the non-replication of a study by Simone Schnall and colleagues. It is worth looking at this controversy in some detail, because it shows how Popper is put to work by the reformers, and how his rules of method and his "friendly-hostile co-operation" are resisted by others.

The original study had produced the kind of attention-grabbing result that the critics see as typical of the discipline's focus on novelty at the expense of rigour: people who feel clean offer milder judgements of moral transgressions. As the title put it: "Cleanliness Reduces the Severity of Moral Judgements" (Schnall, Benton, & Harvey, 2008). Schnall and colleagues described two experiments that had produced this effect. In the first, participants were first

“primed” with cleanliness by having them do a scrambled-sentences task. They had to construct three-word sentences out of sets of four words; in the experimental condition, half the sets contained words related to cleanliness, such as pure, washed, and immaculate. After this task, the participants were asked to rate six morally loaded actions, including putting false information in one’s CV and using a kitten for sexual gratification. In the second experiment, participants were asked to wash their hands after watching a disgusting film clip, and then had to rate the vignettes. In both experiments, participants on average made less severe moral judgements in the experimental condition. In their replication study, Johnson, Cheung, and Donnellan (2014), using the same materials, almost identical procedures, and a much larger sample, failed to find any effect in either experiment.

A rather acrimonious debate ensued.⁹ Schnall, supported by luminaries such as Daniel Gilbert (2014), Daniel Kahneman (n.d.), and Matthew Lieberman (2014), complained about the fact that she had not been allowed to review the final report of the replication (it was a Registered Report; Schnall, 2014b). She also criticised the “crime control mindset” of the reform camp, which, she contended, tends to see every non-replication as an indication of problems in the original study (Schnall, 2014a), and then shames the original researcher over online media. A culture of “replication bullying” had emerged (Schnall, 2014b). One of the replicators, for example, had triumphantly called their attempt “an epic fail” (Donnellan, 2013). Rushing to Schnall’s aid, Gilbert accused the “replication police” of being “shameless little bullies” (D. Gilbert, 2014).

But Schnall also raised questions regarding the value of replication studies per se, how they should be conducted, and by whom. She emphasised, first of all, that experimentation in social psychology is very difficult, much more complicated in fact than in hard sciences like physics. “There are always many reasons for a study to go wrong and everything would have to go right to get the effect” (Schnall, 2014a, para. 29). Therefore, when people without expertise in a particular field of study fail to find the same effect in a replication, we shouldn’t read too much into it. “[B]efore you declare that there definitely is no effect, the burden of proof has to be really high” (Schnall, 2014a, para. 29). More generally, the current emphasis on “direct replications” (Schnall preferred the term “method replications”) is misguided. In a complicated field like social psychology, we shouldn’t expect a particular experiment to produce the same effect every time, even when the experiment is done by experts. Human social behaviour is extremely sensitive to variations in the social and cultural context. Schnall noted, for example, that the participants in the replication study (students at an American university) had evaluated the vignettes much more negatively than the English students in the original study, leading to a ceiling effect in the dependent variable. The failed replication was most likely due to the relative moral laxness of English campus culture, she implied. Rather than putting so much weight on a failed direct replication, we should focus on conceptual replications, in which the same theory is tested with a different procedure. In fact, the connection between physical cleanliness and moral judgement had been conceptually replicated many times, Schnall (2014a) insisted.

In response, proponents of direct replications argued that direct replication is a basic requirement in science, and supported this point with references to Popper. Personality psychologist Sanjay Srivastava, for example, argued that “every experimental report comes with its own repeatability theory” (2014a, para. 5) because each methods section

implies that someone who follows the same procedure will get the same results. This implicit mini-theory is eminently falsifiable, as long as we specify what will count as “the same result” and we spell out the requirements of the experiment. Schnall’s argument that it is up to the replicating researchers to acquire the necessary expertise is wrong, according to Srivastava: “The onus is on the original experimenter to be able to tell a competent colleague what is necessary to repeat the experiment” (Srivastava, 2014a, para. 9). And if they cannot, there is no reason to have any confidence in the original result. At this point, Srivastava quoted from Popper’s *The Logic of Scientific Discovery*, which ends with the line: “No serious physicist would offer for publication, as a scientific discovery, any such ‘occult effect,’ as I propose to call it – one for whose reproduction he could give no instructions” (Popper, 2002a, p. 24). Andrew Wilson wrote a scathing reply to Schnall’s and Kahneman’s insistence that the original researcher should always be consulted for a replication:

once you have published some work then it is fair game for replication, failure to replicate, criticism, critique and discussion. ... We don’t need either your permission or your involvement: the only thing we (should) need is your Methods section and if you don’t like this, then stop publishing your results where we can find them. (2014, para. 3)

Srivastava and other proponents of direct replication acknowledged that falsifying a substantial theory is more complicated than testing the mini-theory implied by the methods section of a report. They admitted the problem noted by Quine (1951) that a theory is never tested in isolation but always in combination with a number of background assumptions, and endorsed Lakatos’s amendments to Popper’s methodology (Srivastava, 2014a). It all starts, however, with establishing the reproducibility of the phenomenon itself, and to this end we must do direct replications: “Only by such repetitions can we convince ourselves that we are not dealing with a mere isolated ‘coincidence’, but with events which, on account of their regularity and reproducibility, are in principle inter-subjectively testable” (Popper, 1972/2002a, p. 23). Were we to drop this requirement, “we would be doing history rather than science” (Srivastava, 2014a, para. 5).

Similar replication controversies have occurred since the Schnall affair.¹⁰ Non-replications keep the debate about fundamental problems in psychology and their solution alive. The proponents of conceptual replication have begun to support their position with arguments drawn from philosophy of science as well. Chris Crandall and Jeff Sherman invoke Duhem and Quine to claim that the “‘failure’ of an empirical test is always ambiguous” (Crandall & Sherman, 2016, p. 94). Conceptual replications, they say, “disperse this ambiguity, and as a result, can contribute more to theoretical development and scientific advance” (p. 94). In a similar vein, Wolfgang Stroebe has attempted to combine a commitment to falsificationism with a preference for conceptual replications. Direct replications are initially important for the original researcher to establish whether the effect is robust, but, given the context-sensitivity of social psychological phenomena, later non-replications do not tell us much. Instead, the focus should shift to conceptual replications, and meta-analyses can determine whether the theory is corroborated or falsified. Psychological theories are not about actual phenomena, which are variable, but about the stable, universal mechanisms that underlie them. With a reference to

Popper's *Conjectures and Refutations*, Stroebe argued that "theories are not refuted by a single inconsistent finding but by studies that support an alternative theory that has greater empirical content" (2016, p. 143). The reformers respond that conceptual replications are certainly important to refine a theory and develop it further, but direct replications are fundamental: "[I]f a phenomenon is not replicable (i.e., it cannot be consistently observed), it is simply not possible to empirically pursue the other goals of science" (LeBel, Berger, Campbell, & Loving, 2017, pp. 8–9).¹¹

How and where criticism should be delivered also remains a hotly debated issue. In September 2016, Susan Fiske wrote a brief guest column for the APS Observer, in which she attacked the way psychology's reformers were going about critiquing the work of others. A first version, which found its way online, contained terms like "trash-talk" and "methodological terrorism" (Fiske as cited in Gelman, 2016). After predictable outrage from the critics, the text was toned down a bit for publication, but Fiske's point remained the same: the critics ("bullies") are promoting a culture of shaming, harassment, and unrestrained hostility, and it is creating victims. "[C]olleagues at all career stages have reported leaving the field because of what they see as sheer adversarial viciousness" (Fiske, 2016, para. 5). She thought this culture of "uncurated, unfiltered denigration" is "encouraged by the new media (e.g. blogs, Twitter, Facebook)" (para. 2), where everyone can publish criticism, regardless of whether it is valid or appropriate. Fiske advocated a return to traditional fora, "with their rebuttals and letters-to-the-editor subject to editorial oversight and peer review for tone, substance, and legitimacy" (para. 7). In response, some of the critics posted a conciliatory petition "Promoting Open, Critical, Civil, and Inclusive Scientific Discourse in Psychology," which attracted 600 signatures (Coan et al., 2016), but the dominant opinion about Fiske's piece was that she was papering over the real problems in psychology (e.g., Gelman, 2016; Yarkoni, 2016). The "tone debate" shows no sign of abating (Chambers, 2017b; Schwarzkopf, 2017), and increasingly includes issues of diversity (Hamlin, 2017; Ledgerwood, 2017).

Rules and practices

I have argued that the reform movement is guided by a Popperian view of how science should be conducted, both as to its methodological rules (replication and falsification in particular) and as to the general culture of openness, mutual criticism, and collaboration in which it needs to be embedded. That is not to say that all aspects of Popper's work are represented in the discourse of the critics;¹² nor that they always name Popper as the source of their ideas about replication, falsification, and criticism; nor that there aren't non-Popperian or even anti-Popperian elements in their ideas. What I have shown is that the critics are trying to create a scientific practice that corresponds to Popper's philosophy of science in several key aspects, sometimes explicitly referring to and discussing his work and that of like-minded philosophers of science. In that sense they are putting philosophy of science, Popper in particular, to work.

What about the relation between rules and practice in these reforms? As I mentioned in the introduction, Mulkay and Gilbert (1981) based their analysis on the philosophical position that that relation is essentially indeterminate: "[N]o rule can specify completely what is to count as following or not following that rule. The terms of a rule always need

to be interpreted in relation to the variable characteristics of specific situations” (p. 400). This conception of rules stems from a reading of Wittgenstein’s discussion of rule-following in *Philosophical Investigations* (1953). Philosopher Saul Kripke (1982) is its best-known representative. It is a sceptical reading that has been vigorously disputed in philosophy (Baker & Hacker, 1985), as well as in science studies—see, for instance, the discussion between Michael Lynch and David Bloor in Pickering (1992). In the non-sceptical view there is no essential gap between rule and practice, even though on occasion it may be unclear what rule applies and how. Interpretation, translating a rule to a specific situation, is sometimes required, but not “always,” as Mulkay and Gilbert claimed. Following a rule is not the same as interpreting it. Usually we follow a rule blindly; if interpretation were constantly required it would cease to be a rule.

Even if we accept this non-sceptical view of rules, however, we may concede that the meaning of rules is more ambiguous, less self-evident in situations characterised by novelty and uncertainty. Mulkay and Gilbert (1981) argued this is inherently the case in scientific research, where new techniques, instruments, and effects regularly raise the question of how rules should be applied. In such circumstances, disambiguating the relation between methodological rules and scientific practice requires creating “interpretative procedures and social relationships” (p. 404) to make Popper’s methodological rules “effective as constraints” (p. 407). Indeed, such interpretative work is happening in the current debate, where every new failed replication instantly leads to fresh discussion about how this result should be interpreted, and whether it suggests further rules and procedures—the Schnall affair is typical in this regard. But the reformers do more than discuss the interpretation of methodological rules. Where Mulkay and Gilbert speak of a social practice that guides and controls the interpretation of rules, thus bridging the gap between rules and practice, what the reformers are trying to create goes further by giving these rules administrative form (pre-registration, Registered Reports) and creating infrastructure to facilitate an Open Psychology. The rule that scientific statements must be exposed to the risk of falsification is expressed in and as pre-registration, just as openness to criticism is expressed in and as the technology of the Open Science Framework. The rules are institutionalised and materialised, a process I have compared to Popper’s “social engineering.”

At the same time it is clear that this is a work in progress, and the crisis is not over yet. Two issues remain particularly contentious, as I have shown. The first is the question of how scientific debates should be conducted in the context of the new online collaboration and communication platforms. This is a context that is quite different from the one that Popper had in mind when he described his “open society” and its “friendly-hostile co-operation.” It is a technological landscape in which a pre-print (reporting, say, a failed replication of a classic experiment in psychology) can be made available to anyone with an internet connection and announced on Twitter, where it subsequently may be discussed by anyone with an account. Within hours, opinions are formed and exchanged, counter-opinions appear, and soon a debate develops that needs only a slight exaggeration, an unfortunate phrase, or a bad joke to spiral out of hand in the way that so many online conversations do. It is doubtful that a return to the classic mode of peer-reviewed discussion in scientific journals is the solution, as Susan Fiske (2016) thinks, or even likely as an option, but it is clear that academia in general is looking for a new ethics of

academic debate in this online world. This too is a problem of interpretation: what does “friendly-hostile co-operation” mean in this novel context? What does civility look like in a tweet, or in a blog post? Is it civil to comment anonymously? And so on.

Whereas the first issue is a more general academic problem, the second is specific to psychology and concerns the question of whether or not direct replications have a role in social psychology, and if so, when. As we have seen, some social psychologists resist the reforms because they do not think it is reasonable to demand that the results of social psychological experiments always be reproducible by following the description in the methods section of the original report. Social behaviour is too sensitive to the social, cultural, and historical context, and this context too diverse and changeable, to expect this kind of stability. This issue is at once methodological and ontological: do Popper’s rules of method apply in a field with a uniquely difficult object of research? And *is* that object really so unique? Ironically, it is precisely the direct replications and their recurring failure to reproduce earlier results that have provoked such questions about the appropriateness of a methodology in which direct replication is fundamental. Direct replications have been a disruptive force in psychology over the last six years, and in response arguments for a special status of social psychology, such as those of Crandall and Sherman (2016) and Stroebe (2016), mentioned above, have been formulated. In fact, one could argue that the variability they see as typical of social behaviour has been made visible by direct replications.

This ironic effect is a product of the reformers’ emphasis on methodological rules.¹³ As Mulkay and Gilbert (1981) noted, rules are constraints. Some rules, however, are more constraining than others. Conceptual replication leaves the researcher more freedom than direct replication. In social psychology, that freedom has been used over the years in the production of sameness. By working on a high level of abstraction, that of the theory’s constructs, researchers can claim that their studies provide evidence for the same general theory, even though what happens in the experiments is different in each case. Thus, variations in experimental procedures are used in the production of sameness. If this strategy is combined with publication bias against null-results, as has been the case for decades, the production of sameness is facilitated further, since falsifying instances never see the light of day. Alternatively, if experiments do not quite pan out as hypothesised, the theory may be amended by the addition of further variables (perhaps the relation only holds for women, or it requires a minimum level of anxiety). On this basis, further hypotheses are tested,¹⁴ and again only the successful studies are published. The result is a pattern that was noted (more or less simultaneously) by Paul Meehl (1990) and by Michael Billig (1990), namely that theories in social psychology tend to start out as bold statements of straightforward relations between a few variables, but then amass an increasing number of “refinements” until they become so unwieldy that the field simply loses interest. All the while, the illusion of sameness is maintained, because although the theory gets more and more nuanced, incorporating more and more sources of variability, at some level it has remained the same, and it is shielded from falsification. Variability is never interesting in itself, except as a source of publications: research aims at finding the basic, underlying psychological processes that, in conjunction with a changeable context, “produce behaviour.” Nor is variability ever a risk, if only because the publication bias keeps failed studies well out of sight.

In contrast, in a research practice that puts direct replication up front, variability can assert itself, as it were, on its own terms. It is precisely *because* researchers limit themselves to following the same procedure as an earlier experiment, that variability stands out as anomalous. Moreover, it is *because* researchers severely constrain their own freedom with methodological strictures like pre-registration, that variability may appear unimpeded.¹⁵ This is the irony of the reformers' emphasis on methodological rules: strict, direct replication, maximally constraining the experimenter, has produced disruptions, because it allows the object to object, to borrow a phrase from Latour (2000). Social psychology finds itself confronted with an epistemic device, direct replication, that over the last few years has regularly produced interesting differences in the form of non-replications. The ease of making such results public online offsets the publication bias of traditional peer reviewed journals and allows each non-replication to become a spectacle on social media, demanding a response from the original researchers and the field of social psychology in general.

Conclusion: Psychology's epistemic project

The latest crisis in psychology has spawned a reform movement that is proposing thorough changes to psychology's epistemic practices, and is creating the sociotechnical conditions for what it sees as a better, more scientific psychology. Indeed, there is now a Society for the Improvement of Psychological Science (<http://improvingpsych.org/>), which meets yearly to discuss ideas for further changes in methods and practices. With its emphasis on hypothesis testing, direct replication, collaboration, and open, critical debate, the reform program is distinctly Popperian, and proposals and arguments are frequently supported with references to Popper and like-minded philosophers of science. Thus, the reformers stay well within the bounds of a rather traditional conception of science. Nonetheless, this effort to put Popper to work is innovative simply because such a strict adherence to these methodological principles has not been attempted before in psychology. Warnings about low power, publication bias, questionable research practices, and replication failures have been sounded before, but it is the first time these issues are addressed with such a comprehensive program of methodological reform.

Moreover, Popper may also turn out to be a source of renewal in spite of himself. As I've argued, direct replications have been a disruptive force in the discipline recently, regularly producing results that are so different from those of the original study that fundamental questions are raised not only about methodological and statistical practices, but increasingly about the object of study itself. Critics of the reform movement have countered its proposals and projects with the argument that people are intensely context-sensitive beings in an extremely variable and changeable environment. As Crandall and Sherman put it, "In matters of social psychology, one can never step in the same river twice" (2016, p. 94). During the previous crisis in psychology, Ken Gergen (1973) argued something very similar, but he drew from this the conclusion that social psychology is a form of history and should let go of its ambition to be a natural science. As yet, this is a step that critics of the reform movement are unwilling to take. Behind the variable and diverse behaviour they still presume to lie a stable, universal cognitive

mechanism, which will ultimately be described by theories that are built and tested by doing conceptual replications. This makes social psychology a highly theoretical field, however, with no clear relevance for practice.¹⁶ Such a retreat into abstraction may not be to everyone's liking, and as non-replications keep piling up, putting the robustness of results in doubt, some may choose to look for different approaches altogether, away from quantitative methods and a search for causal laws. They will ask a question that is largely absent from the current crisis debate: What is psychology good for? And are quantitative methods and experiments always the best way to bring it about?

Declaration of conflicting interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Finkel, Eastwick, and Reis (2015) and Hamlin (2017) recently used the term "evidentiary value movement." I will use "reform movement," "reformers," and "critics."
2. Brian Earp and David Trafimow speak of "a widespread if vague allegiance to Popperian ideals in contemporary scientists" (2015, p. 1).
3. Mulkay and Gilbert are speaking specifically about publication practices here, but use them as an example of a more general point about the link between rules and actions.
4. Whether there "really is" a crisis in psychology is itself contested; Stroebe and Strack (2014), for example, contest it. For the purposes of this article I will define the "reform movement" as those people who believe there is a crisis, and take part in the discussions about it and/or the practical initiatives to solve it. There are no firm boundaries, but, as will be discussed later, reformers connect in various ways, including social media and the infrastructure of the Open Science Framework. There now exists a Society for the Improvement of Psychological Science (<http://improvingpsych.org/>), further institutionalising the movement.
5. An even more prominent topic of discussion is statistics. Statistical flaws, low power in particular, are seen as a major problem in psychology. More fundamentally, the reliance on null hypothesis significance testing is seen by some as a basic flaw.
6. "Although everybody had heard of Popper, very few have actually read him" (Mulkay & Gilbert, 1981, p. 393).
7. This conclusion was contested by D. T. Gilbert, King, Pettigrew, and Wilson (2016).
8. Iso-Ahola (2017) also disputes the necessity of falsification.
9. See, for example, Bohannon (2014) for a brief report.
10. A recent example is the discussion about a large-scale replication effort testing the "facial feedback hypothesis" (Strack, 2017; Wagenmakers et al., 2016).
11. LeBel et al.'s (2017) paper is titled "Falsifiability is not optional." That direct replication is fundamental is also argued by Heino (2017).
12. For example, Popper's discussion of probability in *The Logic of Scientific Discovery* (1972/2002a, 2002b) is mostly ignored. Only Daniel Lakens (2017a) has referred to it.
13. In what follows I build on an argument I presented in Chapter 9 of *Histories of Human Engineering* (Derksen, 2017).

14. At least, that is the way research should proceed. In practice, HARKing is common.
15. See also Srivastava (2014a) about (direct) replication as a route to discovery.
16. Crandall and Sherman (2016), like Stroebe and Strack (2014), acknowledge that direct replications are important in an applied context, where interventions need to reliably produce the same effect. They do not explain why direct replications would be more likely to succeed outside the laboratory.

References

- Baker, G., & Hacker, P. (1985). *Wittgenstein: Rules, grammar and necessity*. Oxford, UK: Basil Blackwell.
- Billig, M. (1990). Rhetoric of social psychology. In J. Shotter & I. Parker (Eds.), *Deconstructing social psychology* (pp. 47–60). London, UK: Routledge.
- Bishop, D. V. (2013, July 26). Why we need pre-registration [Blog post]. Retrieved from <http://deevybee.blogspot.nl/2013/07/why-we-need-pre-registration.html>
- Bohannon, J. (2014). Replication effort provokes praise—And “bullying” charges. *Science*, 344(6186), 788–789. doi: 10.1126/science.344.6186.788
- Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex*, 49(3), 609–610. doi: 10.1016/j.cortex.2012.12.016
- Chambers, C. (2017a). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton, NJ: Princeton University Press.
- Chambers, C. (2017b, August 5). Why I hate the “tone debate” in psychology and you should too [Blog post]. Retrieved from <http://neurochambers.blogspot.com/2017/08/why-i-hate-tone-debate-in-psychology.html>
- Coan, J., Dunham, Y., Durante, K., Finkel, E., Gabriel, S., Giner-Sorolla, R., ... Vazire, S. (2016, September 30). Promoting open, critical, civil, and inclusive scientific discourse in psychology [Blog post]. Retrieved from <http://www.spsp.org/blog/inclusive-scientific-discourse1>
- Crandall, C. S., & Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, 66, 93–99. doi: 10.1016/j.jesp.2015.10.002
- Derksen, M. (2017). The priming saga: The subtle technology of psychological experimentation. In *Histories of human engineering: Tact and technology* (pp. 177–198). Cambridge, UK: Cambridge University Press. doi: 10.1017/9781107414921
- Dienes, Z. (2008). *Understanding psychology as a science*. London, UK: Palgrave Macmillan.
- Donnellan, B. (2013, December 11). Go big or go home – A recent replication attempt [Blog post]. Retrieved from <https://traitstate.wordpress.com/2013/12/11/go-big-or-go-home-a-recent-replication-attempt/> [Blog no longer available].
- Earp, B., & Trafimow, D. (2015, May 19). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, 6. doi: 10.3389/fpsyg.2015.00621
- Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., ... Nosek, B. A. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, 67, 68–82. doi: 10.1016/j.jesp.2015.10.012
- Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science’s aversion to the null. *Perspectives on Psychological Science*, 7(6), 555–561.
- Finkel, E. J., Eastwick, P. W., & Reis, H. T. (2015). Best research practices in psychology: Illustrating epistemological and pragmatic considerations with the case of relationship science. *Journal of Personality and Social Psychology*, 108(2), 275–297. doi: 10.1037/pspi0000007

- Fiske, S. T. (2016, November 1). A call to change science's culture of shaming. *APS Observer*. Retrieved from <http://www.psychologicalscience.org/publications/observer/2016/nov-16/a-call-to-change-sciences-culture-of-shaming.html>
- Gelman, A. (2016, September 21). What has happened down here is the winds have changed [Blog post]. Retrieved from <http://andrewgelman.com/2016/09/21/what-has-happened-down-here-is-the-winds-have-changed/>
- Gergen, K. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, 26(2), 309–320.
- Gilbert, D. [DanTGilbert]. (2014, May 24). Psychology's replication police prove to be shameless little bullies: psychol.cam.ac.uk/cece/blog (corrected link) [Tweet]. Retrieved from <https://twitter.com/dantgilbert/status/470199929626193921?lang=en>
- Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on “Estimating the reproducibility of psychological science.” *Science*, 351(6277), 1037–1037. doi: 10.1126/science.aad7243
- Gilbert, N., & Mulkay, M. J. (1984). *Opening Pandora's box: A sociological analysis of scientists' discourse*. Cambridge, UK: Cambridge University Press.
- Hamlin, J. K. (2017). Is psychology moving in the right direction? An analysis of the evidentiary value movement. *Perspectives on Psychological Science*, 12(4), 690–693. doi: 10.1177/1745691616689062
- Heino, M. T. J. (2017, June 2). Replication is impossible, falsification unnecessary and truth lies in published articles (?) [Blog post]. Retrieved from <https://mattiheino.com/2017/06/02/replication-is-impossible/>
- Iso-Ahola, S. E. (2017, June 2). Reproducibility in psychological science: When do psychological phenomena exist? *Frontiers in Psychology*, 8(879). doi: 10.3389/fpsyg.2017.00879
- Johnson, D. J., Cheung, F., & Donnellan, M. B. (2014). Does cleanliness influence moral judgments? *Social Psychology*, 45(3), 209–215. doi: 10.1027/1864-9335/a000186
- Kahneman, D. (n.d.). Kahneman commentary (Uploaded by B. Garcia). *Scribd*. Retrieved from <http://www.scribd.com/doc/225285909/Kahneman-Commentary>
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2(3), 196–217. doi: 10.1207/s15327957pspr0203_4
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., ... Nosek, B. A. (2014). Investigating variation in replicability: A “many labs” replication project. *Social Psychology*, 45(3), 142–152. doi: 10.1027/1864-9335/a000178
- Kripke, S. A. (1982). *Wittgenstein on rules and private language: An elementary exposition*. Cambridge, MA: Harvard University Press.
- Lakens, D. (2017a, June 19). Verisimilitude, belief, and progress in psychological science [Blog post]. Retrieved from <http://daniellakens.blogspot.com/2017/06/verisimilitude-belief-and-progress-in.html>
- Lakens, D. [lakens]. (2017b, July 9). But just think of the number of years you'll benefit from having read this must-read book. [Tweet]. Retrieved from <https://twitter.com/lakens/status/884081460319322112?s=09>
- Latour, B. (2000). When things strike back: A possible contribution of “science studies” to the social sciences. *British Journal of Sociology*, 51(1), 107–123.
- LeBel, E. P. (2017, May 18). The language of science: A primer [Blog post]. Retrieved from <https://proveyourselfwrong.wordpress.com/2017/05/18/the-language-of-science-a-primer/>
- LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017, January 11). *Falsifiability is not optional*. Retrieved from osf.io/preprints/psyarxiv/dv94b
- LeBel, E. P., & Peters, K. R. (2011). Fearing the future of empirical psychology: Bem's (2011) evidence of psi as a case study of deficiencies in modal research practice. *Review of General Psychology*, 15(4), 371–379. doi: 10.1037/a0025172

- Ledgerwood, A. (2017, January 24). Why the f*ck I waste my time worrying about equality [Blog post]. Retrieved from <http://incurablynuanced.blogspot.com/2017/01/inequality-in-science.html>
- Lieberman, M. D. (2014, August 22). Latitudes of acceptance: A conversation with Matthew D. Lieberman. *Edge*. Retrieved from <http://edge.org/conversation/latitudes-of-acceptance>
- Meehl, P. E. (1990). Appraising and amending theories: The strategy of Lakatosian defense and two principles that warrant it. *Psychological Inquiry*, 1(2), 108–141. doi: 10.1207/s15327965pli0102_1
- Mulkay, M., & Gilbert, G. N. (1981). Putting philosophy to work: Karl Popper's influence on scientific practice. *Philosophy of the Social Sciences*, 11(3), 389–407. doi: 10.1177/004839318101100306
- Neuroskeptic. (2013, April 25). For preregistration in fundamental research [Blog post]. Retrieved from <http://blogs.discovermagazine.com/neuroskeptic/2013/04/25/for-preregistration-in-fundamental-research/>
- Open Science Collaboration. (2015, August 28). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716-1–aac4716-7. doi: 10.1126/science.aac4716
- Pickering, A. (Ed.). (1992). *Science as practice and culture*. Chicago, IL: University of Chicago Press.
- Popper, K. R. (2002a). *The logic of scientific discovery* (2nd ed.). London, UK: Taylor & Francis. (Original work published 1972)
- Popper, K. R. (2002b). *The open society and its enemies* (7th ed.). London, UK: Taylor & Francis.
- Quine, W. V. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60(1), 20–43.
- Schepens, W. [vpro noorderlicht]. (2008, May 16). Bekijk de Noorderlicht-aflevering "Wat is wetenschap?" [Watch the Noorderlicht episode "What is science?"] [Video file]. Retrieved from <https://www.youtube.com/watch?v=uNrNZqtTdvA>
- Schnall, S. (2014a, November 18). Moral intuitions, replication, and the scientific study of human nature. Retrieved from <http://edge.org/conversation/simone-schnall-moral-intuitions-replication-and-the-scientific-study-of-human-nature>
- Schnall, S. (2014b, June 23). Simone Schnall on her experience with a registered replication project [Blog post]. Retrieved from <http://www.spsblog.org/simone-schnall-on-her-experience-with-a-registered-replication-project/>
- Schnall, S., Benton, J., & Harvey, S. (2008). With a clean conscience: Cleanliness reduces the severity of moral judgments. *Psychological Science*, 19(12), 1219–1222. doi: 10.1111/j.1467-9280.2008.02227.x
- Schwarzkopf, S. (2017, August 7). Is open science tone deaf? [Blog post]. Retrieved from <https://neuroneurotic.net/2017/08/07/is-open-science-tone-deaf/>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology. *Psychological Science*, 22(11), 1359–1366. doi: 10.1177/0956797611417632
- Spellman, B. A. (2015). A short (personal) future history of revolution 2.0. *Perspectives on Psychological Science*, 10(6), 886–899. doi: 10.1177/1745691615609918
- Srivastava, S. (2014a, July 1). Some thoughts on replication and falsifiability: Is this a chance to do better? [Blog post]. Retrieved from <https://hardsci.wordpress.com/2014/07/01/some-thoughts-on-replication-and-falsifiability-is-this-a-chance-to-do-better/>
- Srivastava, S. (2014b, November 19). Popper on direct replication, tacit knowledge, and theory construction [Blog post]. Retrieved from <https://hardsci.wordpress.com/2014/11/19/popper-on-direct-replication-tacit-knowledge-and-theory-construction/>
- Srivastava, S. (2015, March 12). An open review of Many Labs 3: Much to learn [Blog post]. Retrieved from <https://hardsci.wordpress.com/2015/03/12/an-open-review-of-many-labs-3-much-to-learn/>

- Srivastava, S. [hardsci]. (2017, July 13). I asked my lab what do you want to do this summer. They said let's nerd out of philosophy and statistics together. I love my lab [Tweet]. Retrieved from <https://twitter.com/hardsci/status/885542833771298816>
- Strack, F. (2017, May 16). From data to truth in psychological science. A personal perspective. *Frontiers in Psychology*, 8(702). doi: 10.3389/fpsyg.2017.00702
- Stroebe, W. (2016). Are most published social psychological findings false? *Journal of Experimental Social Psychology*, 66, 134–144. doi: 10.1016/j.jesp.2015.09.017
- Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, 9(1), 59–71. doi: 10.1177/1745691613514450
- Wagenmakers, E.-J., Beek, T., Dijkhoff, L., Gronau, Q. F., Acosta, A., Adams, R. B., Jr., ... Zwaan, R. A. (2016). Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11, 917–928. doi: 10.1177/1745691616674458
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7(6), 632–638. doi: 10.1177/1745691612463078
- Wilson, A. (2014, May 26). Psychology's real replication problem: Our methods sections [Blog post]. Retrieved from <http://psychsciencenotes.blogspot.co.uk/2014/05/psychologys-real-replication-problem.html>
- Wittgenstein, L. (1953). *Philosophische untersuchungen* [Philosophical investigations] (G. E. M. Anscombe, Trans.). Oxford, UK: Basil Blackwell.
- Yarkoni, T. (2016, October 1). There is no “tone” problem in psychology [Blog post]. Retrieved from <https://www.talyarkoni.org/blog/2016/10/01/there-is-no-tone-problem-in-psychology>

Author biography

Maarten Derksen is Assistant Professor of Theory and History of Psychology at the University of Groningen, Netherlands. He wrote about the dynamic between control and resistance in his recent book *Histories of Human Engineering: Tact and Technology* (Cambridge, 2017). He is also interested in the historical and philosophical aspects of the latest crisis in psychology.